Available online at www.sciencedirect.com

SCIENCE DIRECT

Journal of International Money and Finance
24 (2005) 944–958

Journal of International Money and Finance

www.elsevier.com/locate/econbase

An evaluation framework for alternative VaR-models

Dennis Bams [a,1], Thorsten Lehnert [a], Christian C.P. Wolff [a,b,*]

[a] *Limburg Institute of Financial Economics (LIFE), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands*
[b] *CEPR, UK*

Abstract

In this paper we investigate the ability of different models to produce useful VaR-estimates for exchange rate positions. Our analysis shows that it is important to take into account parameter uncertainty, since this leads to uncertainty in the predicted VaR. We make this uncertainty in the VaR explicit by means of simulation. Our empirical results suggest that more sophisticated tail-modeling approaches come at the cost of more uncertainty about the VaR-estimate itself. We show how to adjust VaR calculations in order to take the parameter uncertainty into account. This is accomplished through a data-driven method to deliver not just a point estimate of the VaR, but a *region*.
© 2005 Elsevier Ltd. All rights reserved.

JEL classification: C22; C52; G10

Keywords: Value-at-Risk; Financial time series; Exchange rate positions; GARCH; Estimation risk; Fat-tail distributions

* Corresponding author. Limburg Institute of Financial Economics (LIFE), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31 43 388 3556; fax: +31 43 388 4875.
E-mail address: c.wolff@berfin.unimaas.nl (C.C.P. Wolff).
[1] Other address: ING Group, Amsterdam, The Netherlands.

0261-5606/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jimonfin.2005.05.004

1. Introduction

A practical risk management tool for financial institutions nowadays is Value-at-Risk (VaR). It is prescribed by the Basle Committee to report to the (international) supervisor and is also used as an internal management tool to see, among other things, whether traders remain within their limits. VaR is defined as an amount lost on a portfolio with a given probability over a fixed number of days. The confidence level reflects 'extreme market conditions' with a probability of, for example, 99%. In other words, in only 1% of the cases will we lose more than the reported VaR of our portfolio.

For modeling financial returns and determining the downside risk of financial positions, exact knowledge of the frequency of extreme events is crucial. In the current study, we propose a data-driven method, based on the uncertainty of the estimated parameters of a distribution function for financial returns, to deliver not just a point estimate of a VaR quantile, but a *region*. Our focus is on extreme deviations in exchange rate positions. Different models have been proposed in the literature to model such extreme exchange rate returns. These models often include distribution functions that allow for fat tails, like the Student-t distribution, and time-varying volatility specifications, such as the GARCH specification. In the context of the latter type of specification, a large unexpected shock leads to an increase in the level of consecutive volatilities. The choice of an adequate distribution function is an important one: if a particular distribution does not allow for an empirical phenomenon which is present in the data, then the accuracy of VaR predictions will correspondingly suffer.

The class of GARCH models has been very successful in modeling significant volatility clustering and other properties of financial returns data (Bollerslev et al., 1994). Various studies show the improvement in VaR estimations associated with GARCH models with disturbances that follow fat-tailed distributions. Pownall and Koedijk (1999) extend the standard J.P. Morgan RiskMetrics™ approach and allow for conditional leptokurtosis using a tail estimator to determine the degrees of freedom of the Student-t distribution. Mittnik and Paolella (2000) and Mittnik et al. (2000) demonstrate that more general GARCH structures and skewed fat-tailed distributions (a skewed Student-t or a skewed stable distribution) improve the precision of out-of-sample VaR calculations.

These findings may encourage the adoption of sophisticated distribution functions, which embody different fat-tail characteristics, and of complicated models to allow for heteroskedasticity. In this paper we argue that this is not necessarily the preferred approach. In VaR-applications, it is not only the distribution functions that play an important role, but also the parameter values of these distribution functions. Parameters are usually estimated based upon historical data. When a particular phenomenon is not present in the historical data, the parameters of the distribution function that are intended to account for the phenomenon are estimated with considerable uncertainty, as reflected by the standard errors of the parameter estimates.

Uncertainty in the parameter estimates leads to uncertainty in the underlying distribution function and, hence, uncertainty in the implied VaR. In the empirical part of this paper we will report VaR-estimates for different models. The VaR-estimates will include an average or expected VaR, a median VaR, and confidence intervals that reflect the uncertainty around the expected VaR. We will show that it is important to take estimation risk into account. Estimation risk refers to the fact that point estimates of parameters, resulting from an estimation procedure, do not necessarily correspond to the underlying true parameters. There is still uncertainty about these true values. Ignoring it may lead to an over- or an underestimation of the actual VaR. Barberis (2000) also takes parameter uncertainty into account when considering predictability in future stock returns. The trade-off between the location and the precision of reported VaR provides a first yardstick to evaluate the adequacy of VaR-models.

In this paper we focus on out-of-sample selection for VaR-models.[2] There are at least two reasons why an in-sample selection method does not lead to the optimal VaR-model. First, this method is not applicable for models that are non-nested. Second, the fact that a particular model fits historical data best does not mean that it also provides the best VaR-forecast. In the out-of-sample approach we split the data sample into two parts. The first part is used to estimate the models. The second part is used to compare the forecasted VaRs with corresponding realizations. Repeating this comparison for many different sub-samples allows us to evaluate whether the realizations are consistent with the predicted VaR at the given confidence level.

The aim of this paper is to provide an empirical selection approach to arrive at the most suitable VaR-model for a given dataset. The purpose of such a model is extreme loss forecasting. In our view, such an approach should deal with uncertainty in the reported VaR that stems from parameter uncertainty. The out-of-sample selection method described in the previous paragraph is in our view suitable to arrive at the appropriate VaR-model. First, it focuses on the purpose of the model. Second, it allows for a comparison of alternative models, and, third, it can take into account the uncertainty in the forecasted VaR. Furthermore, we propose an adjustment to the 'best estimate'-predicted VaR to account for parameter uncertainty. In order to illustrate our approach, we will focus on VaR-estimates in the context of exchange rate positions from the point of view of a currency trader.

In the next section we set up the econometric framework, in Section 3 we describe the data, Section 4 provides the empirical results and Section 5 concludes.

2. Econometric framework

Usually, financial time series are not modeled in terms of prices but in terms of returns. In the empirical part of the paper we deal with exchange rates returns. We

[2] The so-called backtesting approach is recommended by the Basle Committee on Banking Supervision (1996) and also used in Pownall and Koedijk (1999) and Mittnik and Paolella (2000).

consider two types or families of models to describe the return on exchange rates. Also, restricted versions of these two types are taken into consideration. The first model is an AR(1)-GARCH(1,1) with (scaled) Student-t innovations (Bollerslev, 1986; Bollerslev et al., 1992). This is a popular description of financial return series, which has been advocated by Mittnik and Paolella (2000) and McNeil and Frey (2000), among others. The model reads:

$$r_t = \mu + \rho(r_{t-1} - \mu) + \varepsilon_t \tag{1}$$

$$\varepsilon_t \sim t\left(0, \sigma_t^2, \theta\right) \tag{2}$$

$$\sigma_t^2 = \beta_0 + \beta_1 \sigma_{t-1}^2 + \beta_2 \varepsilon_{t-1}^2, \tag{3}$$

where r_t denotes the return at time t, μ is the expected return and ρ is the mean-reversion parameter. The error term ε_t is Student-t distributed with scale parameter σ_t^2 at time t, and degrees of freedom θ. Eq. (3) describes the essence of a GARCH(1,1) process. Special cases of this first type of model are the random walk specification ($\rho = 0$), the normal distribution ($\theta \rightarrow \infty$) and the constant volatility model ($\beta_1 = \beta_2 = 0$).

The second family of models that are taken into consideration is the family of stable Paretian distributions. Stable Paretian distributions have a long history in modeling financial returns; see, e.g., Mandelbrot (1963), and Fama and Roll (1968). The encompassing model for this stable family reads:

$$r_t = \mu + \rho(r_{t-1} - \mu) + \eta_t \tag{4}$$

$$\eta_t \sim S(0, c_t, \alpha) \tag{5}$$

$$c_t^\delta = \gamma_0 + \gamma_1 c_{t-1}^\delta + \gamma_2 |\eta_{t-1}|^\delta. \tag{6}$$

The error terms, η_t, follow a stable Paretian distribution with tail index α. The scale parameter is denoted by c_t. Note that the tail behavior, as well as the shape in the center of the distribution, is determined by the parameter α, which is referred to as the characteristic exponent. Heteroskedasticity is modeled in Eq. (6) by a power-GARCH model, in which the parameter δ is estimated along with the other parameters (Ding et al., 1993). The reason for using this particular specification is that the characteristic exponent α of the stable distribution gives a restriction on the number of moments that exist. Note that $\delta = \alpha$ is the limiting case and is not defined. Mittnik et al. (2000) first derived a closed form expression for a symmetric power-GARCH process with stable innovations. They show that as δ approaches α, the process is no longer covariance stationary. Some GARCH–stable models in the literature are estimated setting δ equal to α and therefore fail to address the correct measure of stationarity (the paper by Liu and Brorsen (1995) is a recent example). In general the power-GARCH process leads to a more flexible specification than the usual GARCH-model in which this parameter is restricted ($\delta=2$). The normal distribution is a special case of the stable distribution, which follows by setting $\alpha=2$.

The stable distribution is not known in analytical form. Accurate numerical values for the density of the stable distributions can be calculated by Fourier-transforming the characteristic function, and evaluating the integral numerically. We used Romberg integration, which allows specification in advance of the tolerated error and, in fact, calculation of the density as precisely as is necessary (see Lambert and Lindsey, 1999). The characteristic function of the stable Paretian distribution is given by Mandelbrot (1963) and Fama and Roll (1968):

$$S(q)=\exp\left\{\mathrm{i}\lambda q - c^{\alpha}|q|^{\alpha}\left[1+\mathrm{i}\beta\,\mathrm{sgn}(q)\tan\left(\frac{1}{2}\pi\alpha\right)\right]\right\} \quad 0<\alpha\leq 2, \tag{7}$$

where (α, β, λ, c) are the parameters that characterize each stable distribution. The parameter $\lambda\in(-\infty,\ \infty)$ is said to be the *location* parameter, $\beta\in(-\infty,\ \infty)$ is the *skewness index*, and $\alpha\in(0,\ 2]$ is the *characteristic exponent*. Remember that c is the *scale* parameter. In the empirical application we estimate parameters for the stable distribution. This second family of distributions allows for more complex behavior in the tails than the first family of distributions, the GARCH–Student-t family.

The two families of models that are presented in Eqs. (1)–(3) and (4)–(6), respectively, are non-nested, so an in-sample approach to check which of the two best describes the data is not a viable way to proceed. By choosing the two families of distributions as given above, we seek to demonstrate the strengths and weaknesses of a complicated VaR-model [Eqs. (4)–(6)] relative to a more standard type of model [Eqs. (1)–(3)]. The same approach is applicable to other types of distributions. A complicated model allows for more complex shapes of the tails and, hence, has the potential advantage of better describing the VaR. At the same time the more complicated model may (or may not) lead to more uncertainty in the parameters and hence in the VaR-estimate itself.

Let p denote the vector of unknown parameters and let $\ln L(p)$ denote the associated log-likelihood. The covariance matrix of parameter estimates follows as:

$$C=-\left(\frac{\partial^2 \ln L(p)}{\partial p \partial p'}\right)^{-1}\Bigg|_{p=\hat{p}}, \tag{8}$$

where $\hat{p}$ denotes the vector of maximum-likelihood point estimates for the unknown parameters. In the out-of-sample algorithm that we propose, this covariance matrix plays a crucial role since it reflects parameter uncertainty. We will use it to quantify the uncertainty in the VaR.

Denote the vector of parameters of the return model by p. The VaR of a position with a confidence level q and for N periods ahead follows as:

$$\mathrm{VaR}_{q,N}(p) = W_0\left\{1 - Q_{1-q}\left(\exp\sum_{t=1}^{N} r_t\right)\right\}, \tag{9}$$

where W_0 is the initial value of the position, and r_t denotes the continuously compounded exchange rate return at time t. Because the parameter values in the underlying return generating process (and the uncertainty therein) have an impact on the VaR, we have made this relationship explicit by writing the VaR as a function of p. Finally, $Q_{1-q}(\cdot)$ denotes the quantile-operator that calculates the $(1-q)$-th quantile of the random variable between parentheses. Usually, an analytical expression for the quantile-operator is not available and we have to rely on simulation instead. A sample path of returns is generated by plugging the parameter estimates $\hat{p}$ into the return model and by sampling from the error term. Consider D samples where $r_1^{(i)}, \ldots, r_N^{(i)}$ denotes the i-th sample path of returns. Let $C^{(i)} = \exp\sum_{t=1}^{N} r_t^{(i)}$ denote the associated cumulative return. An estimate for the $(1-q)$-th quantile follows by first sorting the D cumulative returns in ascending order, and then choosing the $(1-q)D$-th element of this sorted series, denoted as $\hat{Q}_{1-q}$. The associated VaR follows easily from Eq. (9).

Parameter uncertainty may be incorporated by sampling from the parameter distribution. Asymptotic distribution theory leads to the following distribution for the parameter estimates:

$$\hat{p} \sim N(p, C), \tag{10}$$

where $\hat{p}$ are the parameter values that maximize the log-likelihood function, and C denotes the associated covariance matrix of the parameter estimates. In a Bayesian framework we sample from:

$$p \sim N(\hat{p}, C). \tag{11}$$

Consider M samples, which are denoted as $p^{(1)}, \ldots, p^{(M)}$. For all these parameter values we calculate the VaR following the procedure outlined above. This leads to M values for the VaR, denoted as $\mathrm{VaR}_{q,N}(p^{(1)}), \ldots, \mathrm{VaR}_{q,N}(p^{(M)})$. So, instead of arriving at one VaR, we now have an entire *sample* of VaRs. The uncertainty in the VaR may be quantified by calculating the confidence intervals of the VaRs. The expression of the VaR in Eq. (9) shows that both the return distribution and the parameters may be treated as random variables that impact on the size and the uncertainty in the estimated VaR.

To test the adequacy of the return model for VaR-purposes, we propose an out-of-sample analysis in which the data series is split into two parts. Let T_1 denote the length of the first part. In the k-th sub-sample of the first part we use the observations $r_k, \ldots, r_{k+T_1}$ to determine the parameters of the return model, denoted as $\hat{p}_k$, and the associated covariance matrix, denoted as C_k. These suffice to calculate the magnitude of, and the uncertainty in, the VaR from the first part, denoted as $\mathrm{VaR}^k_{q,N}(p_k^{(1)}),\ldots,\ \mathrm{VaR}^k_{q,N}(p_k^{(M)})$. This procedure may be repeated for other sub-samples $k=1,\ldots,\ K$. Associated with the k-th sub-sample is a second part of the data, which covers the observations $r_{T_1+k},\ldots,\ r_T$. From these returns we are able to calculate the *actual* change of the position in a period of length N, denoted as

$$\Delta_N^k = W_{T_1+k-1}\left(1-\exp\sum_{i=1}^{N} r_{T_1+k-1+i}\right). \tag{12}$$

By comparing the predicted VaR with the actual change in the position for a particular sub-sample, we may calculate the number of violations of the predicted VaR. Because of uncertainty in the predicted VaR, the number of violations is different when we use the average predicted VaR instead of using some worst-case VaR. Suppose that for all sub-samples we choose the l-th quantile to represent the VaR for the particular sub-sample. Define the series:

$$V_{q,N}^{k,l} = Q_l\left(\mathrm{VaR}^k_{q,N}(p)\right) \quad k=1,\ldots,\ K, \tag{13}$$

where $Q_l(\cdot)$ denotes the quantile-operator that determines the l-th quantile from the random variable between parentheses. Let the total number of violations associated with the series in Eq. (13) be calculated as:

$$v^l \equiv \sum_{k=1}^{K} I\left\{\Delta_N^k > V_{q,N}^{k,l}\right\}, \tag{14}$$

where $I\{\cdot\}$ is the indicator function. The numbers that result from Eq. (14) allow us to verify below whether a particular quantile of the predicted VaR leads to an appropriate number of violations in our empirical applications. Also, they may lead to an adjustment of the best estimate VaR in order to account for uncertainty in the parameters.

3. Data

Our dataset consists of daily prices of foreign currencies in terms of the US dollar. The currencies include the Deutschmark (DM), the British pound (BP), the Japanese yen (JY) and the Swiss franc (SF). The time span includes the period from January 1986 to September 1999, a total of $T = 3445$ observations. The data are obtained from Datastream. The raw exchange rates are transformed into continuously compounded returns, according to:

$$r_t = 100 \times \ln\left(\frac{\mathrm{ER}_t}{\mathrm{ER}_{t-1}}\right), \tag{16}$$

where ER_t denotes the exchange rate at time t. Summary statistics (not reported) indicate that all currencies' returns exhibit excess kurtosis.

4. Empirical results

In this section we analyze the properties of the alternative VaR-models that are under consideration. We start with an in-sample analysis. Next, VaR-estimates are

Table 1
VaR results Student-t model[a]

N	q (%)		DM	BP	JY	SF
5	99	Mean	4.84	5.20	5.79	5.19
		Median	4.84	5.20	5.79	5.19
		[LB, UB]	[4.48, 5.24]	[4.78, 5.62]	[5.33, 6.25]	[4.78, 5.61]
	95		3.23	3.34	3.69	3.50
			3.23	3.34	3.67	3.49
			[3.02, 3.45]	[3.11, 3.57]	[3.43, 3.95]	[3.26, 3.75]
	90		2.47	2.52	2.76	2.67
			2.46	2.52	2.76	2.66
			[2.30, 2.64]	[2.34, 2.70]	[2.57, 2.95]	[2.48, 2.86]
10	99	Mean	6.67	7.01	7.85	7.27
		Median	6.66	7.01	7.85	7.25
		[LB, UB]	[6.16, 7.16]	[6.44, 7.51]	[7.22, 8.40]	[6.69, 7.78]
	95		4.59	4.72	5.21	4.90
			4.57	4.71	5.18	4.90
			[4.27, 4.95]	[4.39, 5.10]	[4.85, 5.63]	[4.56, 5.29]
	90		3.55	3.66	3.95	3.78
			3.55	3.65	3.95	3.78
			[3.28, 3.83]	[3.37, 3.95]	[3.63, 4.27]	[3.48, 4.08]
20	99	Mean	9.20	9.72	10.85	9.94
		Median	9.20	9.71	10.82	9.93
		[LB, UB]	[8.42, 9.98]	[8.94, 10.60]	[9.98, 11.83]	[9.14, 10.84]
	95		6.47	6.77	7.34	6.98
			6.45	6.77	7.33	6.97
			[5.89, 7.04]	[6.16, 7.38]	[6.67, 8.00]	[6.35, 7.61]
	90		5.07	5.32	5.67	5.43
			5.06	5.31	5.65	5.43
			[4.56, 5.58]	[4.78, 5.86]	[5.10, 6.24]	[4.88, 5.97]

[a] The table reports VaR-results for different horizons (N in days) and at different confidence levels (q). The mean and median VaRs are reported. The uncertainty in the VaR is reflected by the lower bound (LB) and upper bound (UB) of the distribution of simulated VaR-estimates. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

calculated given an initial position of $W_0 = 100$, at different confidence levels and for different time-spans. A measure of the reliability of the reported VaR is presented with the VaR-estimates. Finally, an out-of-sample analysis of the models is carried out, in which the estimated VaR is compared with realized changes in the currency position for different sub-periods. This allows us to evaluate how the alternative models perform in practice.

In line with the literature, likelihood ratio tests (not reported), based on preliminary in-sample estimation results, reject the Student-*t* model in favor of the GARCH(1,1)−Student-*t* model, and reject the GARCH(1,1)−*N* model in favor of the GARCH(1,1)−Student-*t* model. Testing the power-GARCH(1,1)−stable model against the GARCH(1,1)−*N* model results in a preference for the former model. These in-sample model comparisons suggest that the GARCH(1,1)−Student-*t* and the power-GARCH(1,1)−stable models are preferred over restrictive versions of these models.

Table 2
VaR results GARCH(1,1)−*N* model[a]

N	*q* (%)		DM	BP	JY	SF
5	99	Mean	3.67	2.85	4.73	4.13
		Median	3.67	2.85	4.73	4.12
		[LB, UB]	[3.44, 3.90]	[2.67, 3.02]	[4.43, 4.98]	[3.89, 4.36]
	95		2.55	1.99	3.30	2.88
			2.55	1.99	3.29	2.88
			[2.40, 2.72]	[1.85, 2.12]	[3.08, 3.50]	[2.69, 3.06]
	90		1.98	1.55	2.57	2.23
			1.97	1.55	2.56	2.23
			[1.84, 2.52]	[1.44, 1.67]	[2.38, 2.75]	[2.08, 2.40]
10	99	Mean	5.16	4.07	6.60	5.85
		Median	5.16	4.06	6.60	5.85
		[LB, UB]	[4.80, 5.52]	[3.78, 4.35]	[6.14, 7.10]	[5.44, 6.26]
	95		3.58	2.86	4.63	4.06
			3.58	2.86	4.63	4.06
			[3.31, 3.86]	[2.63, 3.09]	[4.26, 5.00]	[3.74, 4.38]
	90		2.78	2.25	3.62	3.18
			2.78	2.24	3.62	3.17
			[2.53, 3.03]	[2.05, 2.46]	[3.29, 3.95]	[2.89, 3.46]
20	99	Mean	7.26	5.88	9.16	8.23
		Median	7.26	5.88	9.15	8.22
		[LB, UB]	[6.73, 7.82]	[5.43, 6.32]	[8.47, 9.85]	[7.61, 8.85]
	95		5.06	4.14	6.41	5.76
			5.05	4.13	6.41	5.76
			[4.62, 5.51]	[3.77, 4.55]	[5.83, 7.05]	[5.24, 6.34]
	90		3.94	3.26	5.02	4.48
			3.93	3.26	5.02	4.47
			[3.51, 4.37]	[2.90, 3.62]	[4.47, 5.57]	[3.99, 4.97]

[a] The table reports VaR-results for different horizons (*N* in days) and at different confidence levels (*q*). The mean and median VaRs are reported. The uncertainty in the VaR is reflected by the lower bound (LB) and upper bound (UB) of the distribution of simulated VaR-estimates. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

Table 3
VaR results GARCH(1,1)–Student-t model[a]

N	q (%)		DM	BP	JY	SF
5	99	Mean	4.77	3.86	6.95	5.29
		Median	4.75	3.84	6.91	5.28
		[LB, UB]	[4.40, 5.27]	[3.55, 4.25]	[6.39, 7.61]	[4.87, 5.82]
	95		3.11	2.51	4.25	3.41
			3.11	2.50	4.25	3.41
			[2.88, 3.32]	[2.33, 2.69]	[3.95, 4.55]	[3.17, 3.65]
	90		2.35	1.89	3.15	2.60
			2.35	1.89	3.14	2.60
			[2.19, 2.52]	[1.76, 2.02]	[2.93, 3.37]	[2.42, 2.78]
10	99	Mean	6.83	5.62	10.05	7.49
		Median	6.80	5.60	10.01	7.48
		[LB, UB]	[6.28, 7.49]	[5.17, 6.13]	[9.24,10.96]	[6.89, 8.16]
	95		4.48	3.65	6.23	4.96
			4.47	3.65	6.21	4.95
			[4.15, 4.86]	[3.39, 3.94]	[5.79, 6.73]	[4.61, 5.36]
	90		3.40	2.82	4.59	3.75
			3.40	2.82	4.59	3.75
			[3.12, 3.68]	[2.57, 3.05]	[4.18, 4.96]	[3.41, 4.06]
20	99	Mean	9.96	8.49	14.98	10.71
		Median	9.91	8.46	14.92	10.68
		[LB, UB]	[8.89,11.19]	[7.55, 9.51]	[13.33,16.78]	[9.53,11.99]
	95		6.55	5.54	9.18	7.08
			6.54	5.53	9.17	7.07
			[5.94, 7.16]	[5.04, 6.05]	[8.35,10.01]	[6.44, 7.72]
	90		4.99	4.19	6.75	5.38
			4.99	4.19	6.74	5.38
			[4.49, 5.51]	[3.77, 4.65]	[6.08, 7.49]	[4.84, 5.97]

[a] The table reports VaR-results for different horizons (N in days) and at different confidence levels (q). The mean and median VaRs are reported. The uncertainty in the VaR is reflected by the lower bound (LB) and upper bound (UB) of the distribution of simulated VaR-estimates. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

Models that provide the best in-sample fit do not necessarily lead to the best VaR-estimates. In Tables 1–4, VaR-estimates are presented for different forecasting horizons ($N=5$, 10, 20 days) and for different confidence levels ($q=90\%$, 95%, 99%). Parameter uncertainty leads to uncertainty in the reported VaR. This uncertainty is quantified by reporting confidence regions for the VaRs within parentheses. The confidence regions are based on lower and upper bounds that result from our simulations. Note that the lower bound (LB) and the upper bound (UB) represent the 2.5% and 97.5% quantiles, respectively. They are not based on standard errors. In general, the VaR is higher at higher confidence levels and also for longer forecasting periods. Comparison of the reported VaRs that result from the different models, shows two effects. First, there is a level effect: because models with stable or Student-t distributed error terms include fatter tail-specifications, the reported VaR is also higher. The VaR may even be overstated, because the slow power-law decay of the fitted stable or Student-t distribution

Table 4
VaR results GARCH(1,1)–stable model[a]

N	q (%)		DM	BP	JY	SF
5	99	Mean	4.90	4.11	6.78	5.19
		Median	4.90	4.09	6.75	5.19
		[LB, UB]	[4.14, 5.72]	[3.49, 4.80]	[5.76, 7.57]	[4.41, 6.07]
	95		2.83	2.11	3.41	3.01
			2.83	2.11	3.40	3.01
			[2.42, 3.18]	[1.81, 2.36]	[2.93, 3.82]	[2.59, 3.37]
	90		2.11	1.66	2.45	2.28
			2.11	1.66	2.45	2.27
			[1.80, 2.35]	[1.41, 1.84]	[2.08, 2.72]	[1.94, 2.53]
10	99	Mean	7.17	6.08	10.04	7.66
		Median	7.15	6.05	10.04	7.65
		[LB, UB]	[6.05, 8.35]	[5.11, 7.05]	[8.43,11.65]	[6.43, 8.89]
	95		4.12	3.25	5.02	4.44
			4.12	3.25	5.02	4.43
			[3.60, 4.59]	[2.83, 3.61]	[4.37, 5.57]	[3.86, 4.93]
	90		3.06	2.43	3.61	3.20
			3.06	2.43	3.61	3.20
			[2.67, 3.41]	[2.11, 2.72]	[3.14, 4.08]	[2.78, 3.58]
20	99	Mean	10.37	8.95	14.91	11.47
		Median	10.34	8.91	14.85	11.43
		[LB, UB]	[8.67,12.23]	[7.52,10.26]	[12.52,17.59]	[9.64,13.53]
	95		5.97	4.85	7.32	6.52
			5.97	4.84	7.31	6.50
			[5.21, 6.63]	[4.22, 5.38]	[6.37, 8.13]	[5.67, 7.24]
	90		4.43	3.58	5.18	4.83
			4.43	3.58	5.18	4.83
			[3.86, 4.93]	[3.11, 4.01]	[4.51, 5.80]	[4.20, 5.41]

[a] The table reports VaR-results for different horizons (N in days) and at different confidence levels (q). The mean and median VaRs are reported. The uncertainty in the VaR is reflected by the lower bound (LB) and upper bound (UB) of the distribution of simulated VaR-estimates. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

implies extreme values never observed in financial data. Second, because of greater uncertainty in the parameter estimates that account for the fat-tail behavior (θ and α), the uncertainty in the reported VaR is also greater for these models. For the power-GARCH(1,1)–stable model, an additional source of uncertainty is the δ-exponent in the volatility model. A huge amount of data are necessary to obtain a precise estimate for the δ-exponent. This leads to more parameter uncertainty and hence to more uncertainty in the associated VaR. The empirical results suggest that more sophisticated tail-modeling approaches come at the cost of more uncertainty about the VaR-estimate itself.

In the last part of the empirical analysis we focus on the out-of-sample behavior of the alternative VaR-models. The parts of the dataset that are used to estimate the alternative VaR-models always have the same length of $T_1 = \frac{1}{2}T = 1722$. Given the parameter estimates, the VaRs are calculated for out-of-sample periods of $N = 5,\ 10,\ 20$ days. This procedure is repeated for $K = 1700$ sub-samples that appear

Table 5
Out-of-sample violations Student-t model[a]

N	Violations	DM	BP	JY	SF
5	Mean	0.8%	0.8%	1.1%	1.4%
	Median	0.8%	0.8%	1.1%	1.4%
	[LB, UB]	[0.4%, 1.8%]	[0.3%, 1.7%]	[0.4%, 1.6%]	[0.2%, 1.8%]
10	Mean	0.7%	0.6%	1.3%	0.8%
	Median	0.7%	0.6%	1.4%	0.8%
	[LB, UB]	[0.5%, 1.6%]	[0.4%, 1.7%]	[1.0%, 2.2%]	[0.3%, 1.9%]
20	Mean	1.2%	1.1%	1.6%	1.3%
	Median	1.3%	1.2%	1.8%	1.3%
	[LB, UB]	[0.6%, 2.1%]	[0.7%, 1.6%]	[0.8%, 3.1%]	[0.4%, 2.1%]

[a] The table reports the percentage of violations of the reported VaR with $q = 99\%$. Violations for different horizons (N in days) and different VaRs (mean, median, lower bound (LB) and upper bound (UB)) are reported. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

as moving windows. Parameter uncertainty is explicitly taken into account: from the different VaR-values that arise because of parameter uncertainty, we report the average VaR, the median VaR, and lower and upper bounds. Again, the lower bound (LB) and the upper bound (UB) represent the 2.5% and 97.5% quantiles, respectively. In Tables 5–8 we report the percentages of violations in the out-of-sample periods.

We start with an analysis of the violations without taking parameter uncertainty into account. This means that we only consider the number of violations of the average reported VaR. The Student-t model and the GARCH(1,1)$-N$ model underestimate the actual VaR on a 99% confidence level for the Japanese yen and the Swiss franc for weekly, biweekly and monthly forecasts, while they provide adequate results for the Deutschmark and the British pound. It is well known that the normal distribution underestimates large events and the result is not surprising in

Table 6
Out-of-sample violations GARCH(1,1)$-N$ model[a]

N	*Violations*	DM	BP	JY	SF
5	Mean	1.2%	1.1%	2.1%	1.6%
	Median	1.2%	1.1%	2.1%	1.6%
	[LB, UB]	[0.8%, 2.1%]	[0.5%, 1.9%]	[1.6%, 3.5%]	[1.3%, 2.5%]
10	Mean	0.9%	0.8%	2.6%	1.8%
	Median	0.9%	0.8%	2.6%	1.8%
	[LB, UB]	[0.4%, 1.4%]	[0.4%, 1.4%]	[2.1%, 3.4%]	[0.8%, 2.8%]
20	Mean	1.2%	1.1%	3.7%	1.5%
	Median	1.2%	1.1%	3.7%	1.5%
	[LB, UB]	[0.5%, 2.1%]	[0.8%, 1.5%]	[2.5%, 5.6%]	[0.9%, 2.6%]

[a] The table reports the percentage of violations of the reported VaR with $q = 99\%$. Violations for different horizons (N in days) and different VaRs (mean, median, lower bound (LB) and upper bound (UB)) are reported. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

Table 7
Out-of-sample violations GARCH(1,1)–Student-t model[a]

N	Violations	DM	BP	JY	SF
5	Mean	0.7%	0.6%	0.9%	0.9%
	Median	0.7%	0.6%	1.0%	0.9%
	[LB, UB]	[0.5%, 1.3%]	[0.4%, 1.1%]	[0.5%, 1.4%]	[0.2%, 1.4%]
10	Mean	0.7%	0.6%	1.1%	0.5%
	Median	0.8%	0.6%	1.1%	0.5%
	[LB, UB]	[0.5%, 1.5%]	[0.5%, 1.5%]	[0.5%, 1.8%]	[0.1%, 1.1%]
20	Mean	1.0%	0.9%	0.0%	0.4%
	Median	1.1%	0.9%	0.2%	0.5%
	[LB, UB]	[0.7%, 2.1%]	[0.6%, 1.7%]	[0.0%, 1.5%]	[0.1%, 1.4%]

[a] The table reports the percentage of violations of the reported VaR with $q = 99\%$. Violations for different horizons (N in days) and different VaRs (mean, median, lower bound (LB) and upper bound (UB)) are reported. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

the case of the Japanese yen and the Swiss franc. The Student-t model gives a better description of the tail fatness in FX return data, but the constant volatility model is not flexible enough to capture the behavior of more volatile heteroskedastic time series. The GARCH(1,1)–Student-t model in general overestimates the actual VaR except for the Japanese yen at weekly ($N = 5$) and biweekly ($N = 10$) horizons. For $N = 20$, it leads to no violation on average. The overestimation of the GARCH(1,1)–Student-t model for four out of five exchange rates is a result of the power-law tails of the Student-t distribution. For exchange rate data the really extreme events are rare. Because of the functional specification, really extreme returns are generated much more frequently than is the case in the historical data. As a result, these extreme returns are the source for the overestimation of the actual VaR, especially for monthly forecasts. The same is true for the power-GARCH(1,1)–stable model and the results are similar. Based on the average

Table 8
Out-of-sample violations power-GARCH(1,1)–stable model[a]

N	Violations	DM	BP	JY	SF
5	Mean	0.6%	0.5%	1.1%	0.9%
	Median	0.7%	0.5%	1.1%	0.9%
	[LB, UB]	[0.1%, 1.9%]	[0.1%, 1.7%]	[0.4%, 2.4%]	[0.2%, 2.2%]
10	Mean	0.3%	0.4%	1.3%	0.5%
	Median	0.3%	0.5%	1.4%	0.6%
	[LB, UB]	[0.1%, 1.8%]	[0.0%, 1.9%]	[0.0%, 2.5%]	[0.1%, 2.2%]
20	Mean	0.4%	0.3%	0.1%	0.4%
	Median	0.5%	0.4%	0.4%	0.6%
	[LB, UB]	[0.1%, 1.9%]	[0.0%, 2.0%]	[0.0%, 3.3%]	[0.1%, 2.2%]

[a] The table reports the percentage of violations of the reported VaR with $q = 99\%$. Violations for different horizons (N in days) and different VaRs (mean, median, lower bound (LB) and upper bound (UB)) are reported. The LB and UB are the 2.5% and 97.5% quantiles, respectively.

predicted VaR, we conclude that the GARCH(1,1)−N model leads to an underestimation of the greatest possible loss in the tail because tail behavior is not modeled adequately, and the unconditional Student-t model underestimates the frequency of large events for more volatile processes. The GARCH(1,1)−Student-t and power-GARCH(1,1)−stable models overestimate the risk in the tail. This is due to the fact that parameter estimates do not adequately represent the behavior in the upper tail due to lack of extreme historical observations.

More parameter uncertainty leads to wider confidence intervals for the reported VaR. In particular the GARCH(1,1)−Student-t and power-GARCH(1,1)−stable models lead to wide confidence intervals. This suggests that it is not useful to adopt models with more complex tail-properties, because it will only lead to more uncertainty in the reported VaR. Parameter uncertainty also explains the over-estimation of the reported VaR in case of the GARCH(1,1)−Student-t. Instead of focusing on the average VaR, we may also give more attention to the upper bound. This has the effect that we acknowledge that tails are less fat than the point estimates would imply. This leads to rejection rates that come closer to the confidence levels of the VaR. In case of the power-GARCH(1,1)−stable model there is so much parameter uncertainty that the upper bound (the 97.5% quantile) of the predicted VaR leads to an underestimation. More sophisticated risk adjustments are required in this case.

5. Concluding remarks

In this paper we have focused on VaR-model selection for exchange rate positions. On the one hand, fat-tail behavior is present in exchange rate returns, but on the other hand, there are very few observations far in the tail. Models that take account of tail behavior are required, since otherwise the reported VaR leads to an underestimation of the risk in the tail. Complex tail models often lead to overestimation of the VaR, because these models assume more probability mass in the tail of the distribution than is actually present. This is due to the fact that very few extreme observations occur and hence tail behavior is measured with relatively great uncertainty. For the GARCH(1,1)−Student-t distribution, taking our adjustment into account leads to rejection rates that are close to the confidence levels of the predicted VaR. This makes the GARCH(1,1)−Student-t distribution an adequate model to correctly assess extreme losses for exchange rate positions. Models with more sophisticated tail behavior lead to more parameter uncertainty, which in turn leads to greater uncertainty in the predicted VaRs.

Acknowledgements

We would like to thank Stefan Mittnik, Franz Palm and Peter Schotman for helpful comments. We are also grateful to Michael Melvin (the Co-Editor) and two anonymous referees for their useful suggestions to improve the paper.

References

Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance 55 (1), 225–264.

Basle Committee on Banking Supervision, 1996. Amendment to the Capital Accord to Incorporate Market Risks. Basle Report No. 24, BIS.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modeling in finance: a selective review of the theory and empirical evidence. Journal of Econometrics 52, 5–59.

Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models. In: Engle, R.F., McFadden, D. (Eds.), Handbook of Econometrics, vol. IV. North-Holland, Amsterdam, pp. 2959–3038.

Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83–106.

Fama, E.F., Roll, R., 1968. Some properties of symmetric stable distributions. Journal of the American Statistical Association 63.

Lambert, P., Lindsey, J.K., 1999. Analysing financial returns by using regression models based on non-symmetric stable distributions. Journal of Applied Statistics 48, 409–424.

Liu, S., Brorsen, B.W., 1995. Maximum likelihood estimation of a GARCH–stable model. Journal of Applied Econometrics 10, 273–285.

Mandelbrot, B., 1963. The variation of certain speculative prices. Journal of Business 36, 394–419.

McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7 (3–4), 271–300.

Mittnik, S., Paolella, M.S., 2000. Conditional density and value-at-risk prediction of Asian currency exchange rates. Journal of Forecasting 19, 313–333.

Mittnik, S., Paolella, M.S., Rachev, S.T., 2000. Diagnosing and treating the fat tails in financial return data. Journal of Empirical Finance 7, 389–416.

Pownall, R.A.J., Koedijk, K.G., 1999. Capturing downside risk in financial markets: the case of the Asian crisis. Journal of International Money and Finance 18, 853–870.